



Data Analysis with Python

Overview

Learn how to analyze data using Python. This course will take you from the basics of Python to exploring many different types of data. You will learn how to prepare data for analysis, perform simple statistical analysis, create meaningful data visualizations, predict future trends from data, and more!

Topics covered:

- 1) Importing Datasets
- 2) Cleaning the Data
- 3) Data frame manipulation
- 4) Summarizing the Data
- 5) Building machine learning Regression models
- 6) Building data pipelines

Data Analysis with Python will be delivered through lecture, lab, and assignments. It includes following parts:

Data Analysis libraries: will learn to use Pandas, Numpy and Scipy libraries to work with a sample dataset. We will introduce you to pandas, an open-source library, and we will use it to load, manipulate, analyze, and visualize cool datasets. Then we will introduce you to another open-source library, scikit-learn, and we will use some of its machine learning algorithms to build smart models and make cool predictions.

Skills You Will Gain

- Predictive Modelling
- Python Programming
- Data Analysis
- Data Visualization (DataViz)
- Model Selection

MIND MATRIX SDN. BHD. [201301001419 (1031256 P)]

Suite 33.01, 33rd Floor, Menara Keck Seng, 203, Jalan Bukit Bintang,
55100 Kuala Lumpur, Malaysia.

Tel: 03-2116 5778 | Fax: 03-2116 5999 | Email: info@mindasys.com



Who Should Enroll?

- IT Managers and Project Managers
- Software Developers and ETL Professionals
- Analytics Professionals
- Those who are wanting to have a career in Python

Course Outline

Part 1: Importing Datasets

- The Problem
- Understanding the Data
- Python Packages for Data Science
- Importing and Exporting Data in Python
- Getting Started Analyzing Data in Python

Part 2: Data Wrangling

- Pre-processing Data in Python
- Dealing with Missing Values in Python
- Data Formatting in Python
- Data Normalization in Python
- Binning in Python
- Turning categorical variables into quantitative variables in Python

Part 3: Exploratory Data Analysis

- Exploratory Data Analysis
- Descriptive Statistics
- GroupBy in Python
- Correlation
- Correlation – Statistics
- Analysis of Variance ANOVA

Part 4: Model Development

- Model Development
- Linear Regression and Multiple Linear Regression
- Model Evaluation using Visualization
- Polynomial Regression and Pipelines
- Measures for In-Sample Evaluation
- Prediction and Decision Making

MIND MATRIX SDN. BHD. [201301001419 (1031256 P)]

Suite 33.01, 33rd Floor, Menara Keck Seng, 203, Jalan Bukit Bintang,
55100 Kuala Lumpur, Malaysia.

Tel: 03-2116 5778 | Fax: 03-2116 5999 | Email: info@mindasys.com



Part 5: Model Evaluation

- Model Evaluation and Refinement
- Overfitting, Underfitting and Model Selection
- Ridge Regression
- Grid Search

MIND MATRIX SDN. BHD. [201301001419 (1031256 P)]

Suite 33.01, 33rd Floor, Menara Keck Seng, 203, Jalan Bukit Bintang,
55100 Kuala Lumpur, Malaysia.

Tel: 03-2116 5778 | Fax: 03-2116 5999 | Email: info@mindasys.com

